

UC San Diego

UC San Diego Previously Published Works

Title

A Bayesian approach for statistical-physical bulk parameterization of rain microphysics. Part II: Idealized Markov chain Monte Carlo experiments

Permalink

<https://escholarship.org/uc/item/9hr0q7zf>

Journal

Journal of the Atmospheric Sciences, 77(3)

ISSN

0022-4928

Authors

van Lier-Walqui, M
Morrison, H
Kumjian, MR
[et al.](#)

Publication Date

2020-03-01

DOI

10.1175/JAS-D-19-0071.1

Peer reviewed

A Bayesian Approach for Statistical–Physical Bulk Parameterization of Rain Microphysics. Part II: Idealized Markov Chain Monte Carlo Experiments

MARCUS VAN LIER-WALQUI

*Center for Climate Systems Research, Columbia University, and NASA Goddard Institute for Space Studies,
New York, New York*

HUGH MORRISON

National Center for Atmospheric Research, Boulder, Colorado

MATTHEW R. KUMJIAN AND KARLY J. REIMEL

Department of Meteorology and Atmospheric Science, The Pennsylvania State University, University Park, Pennsylvania

OLIVIER P. PRAT

North Carolina Institute for Climate Studies, North Carolina State University, Asheville, North Carolina

SPENCER LUNDERMAN AND MATTHIAS MORZFELD

Department of Mathematics, The University of Arizona, Tucson, Arizona

(Manuscript received 21 March 2019, in final form 4 September 2019)

ABSTRACT

Observationally informed development of a new framework for bulk rain microphysics, the Bayesian Observationally Constrained Statistical–Physical Scheme (BOSS; described in Part I of this study), is demonstrated. This scheme’s development is motivated by large uncertainties in cloud and weather simulations associated with approximations and assumptions in existing microphysics schemes. Here, a proof-of-concept study is presented using a Markov chain Monte Carlo sampling algorithm with BOSS to probabilistically estimate microphysical process rates and parameters directly from a set of synthetically generated rain observations. The framework utilized is an idealized steady-state one-dimensional column rainshaft model with specified column-top rain properties and a fixed thermodynamical profile. Different configurations of BOSS—flexibility being a key feature of this approach—are constrained via synthetic observations generated from a traditional three-moment bulk microphysics scheme. The ability to retrieve correct parameter values when the true parameter values are known is illustrated. For cases when there is no set of true parameter values, the accuracy of configurations of BOSS that have different levels of complexity is compared. It is found that addition of the sixth moment as a prognostic variable improves prediction of the third moment (proportional to bulk rain mass) and rain rate. In contrast, increasing process rate formulation complexity by adding more power terms has little benefit—a result that is explained using further-idealized experiments. BOSS rainshaft simulations are shown to well estimate the true process rates from constraint by bulk rain observations, with the additional benefit of rigorously quantified uncertainty of these estimates.

1. Introduction

The parameterization of cloud and precipitation microphysics is beset by a number of fundamental difficulties. Because there is a limited theoretical understanding

of many microphysical processes, development of model parameterizations must rely heavily on cloud and precipitation observations. Unfortunately, these observations generally do not provide direct information on the microphysical process rates themselves, and so model developers must consider measurements of bulk quantities such as cloud water content that evolve through the net effects of several interacting microphysical

Corresponding author: Marcus van Lier-Walqui, mv2525@columbia.edu

and dynamical processes. As argued in [Morrison et al. \(2020, hereafter Part I\)](#), the incorporation of such observational information in a microphysics scheme should be posed as a probabilistic problem, with Bayes's theorem defining the solution. Use of Bayesian parameter estimation has the important advantage of quantifying uncertainty within the microphysics scheme in question.

As opposed to Bayesian approaches to state estimation employed in, for example, data assimilation for numerical weather forecasts, parameter estimation and related uncertainty quantification problems are often characterized by strong nonlinearities and relatively low problem dimensionality (order <100 unknowns to estimate)—for these problems Markov chain Monte Carlo (MCMC) samplers are well suited to the task of estimating the parameter probability density function (PDF) owing to their generality and lack of assumptions on the underlying form of the PDF or linearity of model sensitivity to parameter perturbation.

A simplified statistical–physical microphysical modeling approach was proposed in [Part I](#) that facilitates the constraint of scheme parameters and estimation of parametric and structural uncertainty directly from observations using Bayesian methods. This formed the basis for the Bayesian Observationally Constrained Statistical–Physical Scheme (BOSS) detailed in [Part I](#). In short, BOSS is a bulk microphysics scheme (currently liquid only) constructed with minimal assumptions about the parameterized form for the drop size distribution (DSD) or functional forms for the microphysical process rates. In BOSS, process rates are generalized as multivariate power expressions of some set of prognostic DSD moments, where a moment of order k is defined as

$$M_k = \int_{D_{\min}}^{D_{\max}} D^k (\partial N / \partial D) dD, \quad (1)$$

with $\partial N / \partial D$ denoting the DSD. Owing to this unique mathematical structure, BOSS can be easily constructed to evolve any number or choice of prognostic DSD moments, and can also have any number of generalized power expression terms to represent the microphysical process rates. Varying these choices and setting the value of power expression parameters defines the scheme's structure. The complexity of this structure is determined by the number of prognostic moments, and number of terms in the process rate formulations. By contrast, existing bulk schemes typically have a fixed choice of prognostic moments, an assumed DSD form, and have an inflexible form for the process rate calculations. Thus, existing schemes have a limited set of parameters whose values can be

systematically perturbed, which in general does not encompass all sources of physical uncertainty.

Our ultimate goal is to use observations to inform BOSS probabilistically, determining the nature and complexity of its physical sensitivities using the wealth of microphysical information present in, for example, polarimetric radar observations, together with the limited a priori microphysical insights that exist. As a proof of concept, we here show that this is feasible within an idealized modeling framework. This should be considered a minimum requirement for a scheme like BOSS, as use of real observations to constrain BOSS will unavoidably introduce other uncertainties (observational noise/calibration, atmospheric-state uncertainty, forward simulator uncertainty, etc.) These difficulties, and approaches to mitigate them, will be left to future work.

There are numerous potential applications for a full account of structural and parametric uncertainty in microphysics, in particular within the context of data assimilation and probabilistic forecasts. For example, multimodel and multiphysics ensembles are commonly employed in these contexts in an attempt to span the structural uncertainties associated with discrete microphysics parameterization choices ([Berner et al. 2011, 2015](#); [Jankov et al. 2017](#)). These multimodel ensembles help produce forecasts with sufficient spread to characterize model forecast error, but have no clearly defined probabilistic meaning. An ensemble of BOSS simulations should be capable of spanning structural and parametric uncertainty in a unified, consistent way that is informed by observations and also probabilistically reflects both lack of observational constraints on microphysical knowledge and the limited physical insights that do exist. It is also possible that BOSS may be relevant for development of stochastic microphysics schemes, but we do not explore such issues here.

A crucial goal of BOSS is that process-level understanding is gained via observational constraint—the statistical–physical approach we have chosen renders all BOSS parameters physically interpretable, and in the ideal limit of a perfect model and comprehensive observations, we would expect that the representation of microphysical processes within BOSS would approach the “true” process rates of nature. In this sense, constraint of BOSS can be viewed in the context of microphysical “fingerprinting” ([Kumjian and Prat 2014](#); [Moiseev et al. 2015, 2017](#)) or microphysical process retrievals ([Williams 2016](#); [Tridon et al. 2017](#)), with the additional benefit of robust estimation of uncertainties arising from a priori assumptions and observational uncertainties.

In the current paper, Part II of this series, we consider three configurations of BOSS with varying degrees of complexity and explore synthetic constraint of the scheme within one-dimensional idealized rainshaft simulations. Our goals are to show that BOSS parameters can be accurately constrained within such a framework using MCMC, and to compare configurations of BOSS with varying degrees of structural complexity. Further, we analyze the structure of uncertainty in BOSS, including correlation between parameters, uncertainty in modeled microphysical processes, and forecast uncertainty that results from microphysics uncertainty.

Idealized rainshaft simulations are used here because we limit our initial development of BOSS to rain microphysical processes (evaporation, collision-coalescence, breakup, and sedimentation) owing to the additional complexities of ice microphysics, including well-known uncertainties in scattering characteristics of ice- and mixed-phase hydrometeors (Schrom and Kumjian 2018). We also for now ignore nonprecipitating cloud microphysics (drop activation, autoconversion, etc.)—these processes will be addressed in subsequent work. A one-dimensional steady-state rainshaft is chosen so that microphysical sensitivities are isolated from dynamical and thermodynamical feedbacks that can obscure their signal (Morrison and Grabowski 2007; Posselt and Vukicevic 2010; Shipway and Hill 2012). Throughout this paper, we refer to the one-dimensional rainshaft simulation as the “model,” the microphysics scheme used by the model as the “scheme,” and the specific setup for the BOSS scheme (e.g., number and choice of prognostic moments) as the “configuration.”

The primary purpose of this paper is demonstrating how BOSS can be constrained using observations and an MCMC sampler. Of particular interest are the following questions:

- (i) Can idealized observations of prognostic variables be used to unambiguously retrieve the true value of model parameters when the true values are known?
- (ii) Do BOSS parameter PDFs display non-Gaussianity (indicative of nonlinear model response to parameter perturbations), or correlations in uncertainty between parameters (which suggest compensating errors)?
- (iii) Can BOSS match the performance of a traditional bulk parameterization scheme, especially with respect to the representation of microphysical processes? How does this vary with the number of prognostic moments?
- (iv) How does varying BOSS complexity (e.g., the number of terms in process rate power expressions) affect estimation of BOSS parameters and model forecast fidelity?

Question (i) will be answered by performing “known truth” experiments where observations are generated using the same model, scheme, and configuration that are used for parameter estimation. Questions (ii) and (iii) will be addressed by generating synthetic observations using a traditional three-moment bulk microphysics scheme (MORR; see section 3a in Part I for a description), and performing Bayesian parameter estimations with various configurations of BOSS. Question (iv) will be addressed by considering a known-truth abstraction of the BOSS parameterization framework.

The paper is organized as follows. Section 2 gives a description of the methodology, including an overview of BOSS and the MCMC sampling. Section 3 details the experimental design of each test using MCMC and synthetic observations to constrain BOSS. Section 4 presents results and brief discussion. A summary and conclusions are provided in section 5.

2. Methods

a. Overview of BOSS

A detailed description of BOSS is given in Part I, with a brief overview in this subsection. BOSS is based on a simple yet flexible statistical-physical set of equations that captures key rain microphysical processes without assuming a DSD functional form. Process rates for scheme prognostic moments are formulated as generalized power series functions of those prognostic moments. The N -moment normalization of Morrison et al. (2019) allows us to formulate the evolution of a moment by a process rate (e.g., evaporation’s effect on mass M_3) as depending on the N prognostic moments using N parameters per power-law term—one power-law coefficient (a) and $N - 1$ exponent terms (β or δ). Briefly, single-drop processes (i.e., processes acting on individual drops—condensation/evaporation and sedimentation) affecting prognostic moment M_k , where $k \in \{p_1, p_2\}$, in two-moment configurations of BOSS are governed by

$$\frac{dM_k}{dt} = F(T, p, q) \sum_l a_{l,k} M_{p_1}^{1-\beta_{l,k}} M_{p_2}^{\beta_{l,k}}, \quad (2)$$

where $F(T, p, q)$ describes the thermodynamic dependence of the process rate, and the sum over l defines the number of power-law terms used. Equations for drop-drop interaction processes (e.g., collision-coalescence, collisional breakup) are similar, except with the “1” in the exponent replaced with a “2” as follows:

$$\frac{dM_k}{dt} = F(T, p, q) \sum_l a_{l,k} M_{p_1}^{2-\delta_{l,k}} M_{p_2}^{\delta_{l,k}}. \quad (3)$$

Three-moment BOSS process rates depend on three prognostic moments M_k , where $k \in \{p_1, p_2, p_3\}$, with two exponent parameters instead of one:

$$\frac{dM_k}{dt} = F(T, p, q) \sum_l a_{l,k} M_{p_1}^{1-\beta_{1,l,k}} M_{p_2}^{\beta_{1,l,k}-\beta_{2,l,k}} M_{p_3}^{\beta_{2,l,k}} \quad (4)$$

for single-drop processes and

$$\frac{dM_k}{dt} = F(T, p, q) \sum_l a_{l,k} M_{p_1}^{2-\delta_{1,l,k}} M_{p_2}^{\delta_{1,l,k}-\delta_{2,l,k}} M_{p_3}^{\delta_{2,l,k}} \quad (5)$$

for drop-drop interaction processes.

The focus of this work is on rain microphysics. As such, the following microphysical processes are included: condensation/evaporation, collision-coalescence, breakup, and sedimentation. BOSS is flexible and can utilize any combination of prognostic moments and any number of terms in the process rate power expressions; these choices can be tailored depending on the observational constraint and Bayesian methodology used. Herein we utilize a two-moment configuration of BOSS prognosing M_0 and M_3 (M0-M3-B) or a three-moment configuration prognosing M_0 , M_3 , and M_6 (M0-M3-M6-B), similar to the tests in [Part I](#) (here M_k is the k th moment of the DSD). We also test the impacts of including either one or two terms for the process rate power expressions for evaporation in two-moment BOSS (M0-M3-B2). As described in the introduction, this allows systematic testing of the effects of the number of parameters and scheme complexity.

We constrain BOSS with synthetic observations of M_0 and M_3 (and M_6 additionally for M0-M3-M6-B) at 15 vertical levels for an ensemble of rainshaft simulations, as well as with synthetic observations of moment fluxes at the surface (M_3 flux is proportional to rain rate) as might be obtained from a disdrometer. The synthetic observations are produced using a steady-state rainshaft model employing a three-moment configuration of [Morrison et al. \(2009\)](#) (MORR; see [Part I](#) for details). We note that MORR is an appealing choice of “truth” for this first test of BOSS constraint, as the functional form of its process rate formulations are in some cases power laws consistent with BOSS (e.g., for evaporation and sedimentation), and in other cases different functions (e.g., exponential for coalescence and breakup; see [Part I](#) appendix for more detail). Three-moment MORR is chosen instead of two moment so as to investigate the case where the model (e.g., M0-M3-B) has less structural complexity compared with the “true” model.

A practical consideration for using more realistic (e.g., radar) observations is that, owing to the lack of a specified DSD functional form in BOSS, many observational

quantities cannot be easily forward simulated from BOSS prognostic DSD moments. Instead, to perform such forward simulation, observable variables would have to be related to prognostic moments in some other way (e.g., statistically). The recently developed moment-based operator of [Kumjian et al. \(2019\)](#), for example, uses a database of hundreds of millions of observed and bin-simulated DSDs to draw statistical relationships between DSD moments and polarimetric radar variables. This approach has the additional benefit of producing estimates of forward-simulator uncertainty. Testing of such methods with BOSS will be addressed in subsequent research and is outside of the scope of the current work, and we instead limit ourselves in this initial proof-of-concept work to the idealized case where model prognostic DSD moments are directly observed—a condition seldom met with real observations.

Both the BOSS simulations and synthetic observations are based on a steady-state rainshaft model with a vertical depth of 2 km and 80 equally spaced levels. The model is initialized with a constant thermodynamic profile, and a fixed DSD at column top (e.g., fixed values of M_0 , M_3 , M_6)—see [Part I](#) for more details. Relative humidity (RH) conditions are drawn randomly from an even sampling of raining and nonraining days from in situ measurements by the U.S. Climate Reference Network ([Diamond et al. 2013](#); [Leeper et al. 2015](#)). Column-top DSDs are drawn randomly from the database described in [Morrison et al. \(2019\)](#): a distribution of M_0 , M_3 , and M_6 values created by an equal mixture of bin model and disdrometer DSD datasets, with the additional constraint that the reflectivity distribution of this dataset is uniform between 10 and 45 dBZ. The choice of boundary condition distributions affects the constraint of BOSS; for example, tests using a climatological distribution of rain rates (not shown) resulted in weaker constraint of coalescence/breakup parameters owing to the reduced effect of these processes in the predominantly lower-rain-rate cases. The DSD database and the choices made in its construction are described in [Morrison et al. \(2019\)](#).

[Figure 1](#) illustrates schematically how observational constraint of BOSS is performed. Initial conditions are drawn from climatology, disdrometers, and bin model simulations (blue); these initialize both the model using MORR that generates the synthetic observations, as well as the model using BOSS (green). The MCMC algorithm uses the synthetic observations to constrain sampling of BOSS parameters such that the distribution of sampled points are drawn from the posterior parameter PDF $P(\mathbf{x}|\mathbf{y}, \mathcal{M})$ (black; see the following section and [\(6\)](#)—this is mediated by the observational uncertainty

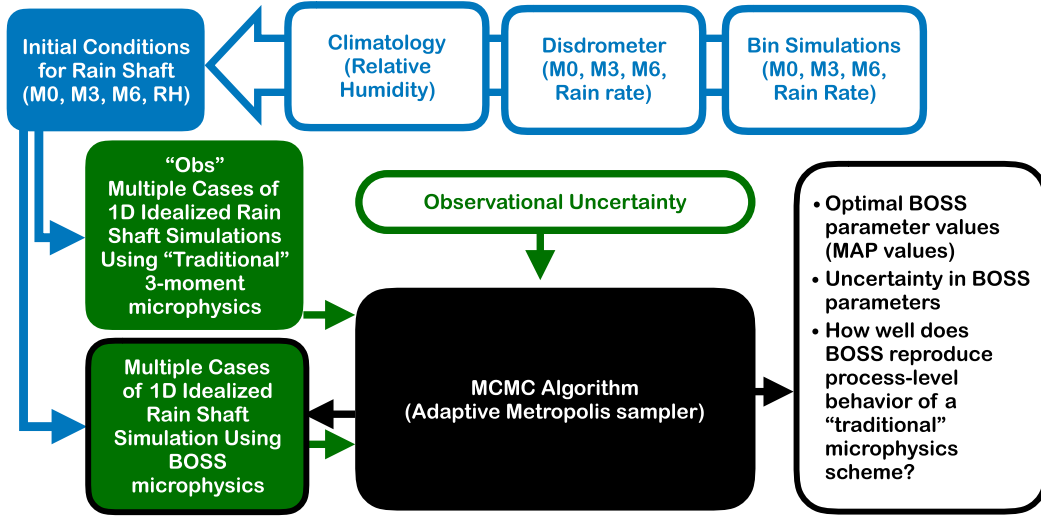


FIG. 1. Schematic of the approach to estimate BOSS parameters in this study.

estimates (green outline), which inform the mismatch between the observations and model output.

b. MCMC

We employ the Adaptive Metropolis (AM; Haario et al. 2001) sampler, an MCMC-like algorithm, to estimate the multivariate posterior parameter PDF, which also provides estimates of maximum a posteriori (MAP; i.e., the most probable) parameter values and attendant uncertainty. MCMC methods employ a modified random walk (chain) through the multidimensional parameter space, with proposal steps in the random walk either accepted or rejected depending on the ratio of posterior probabilities of the previous parameter values and the proposed parameter values. The behavior of AM is similar to that of the basic Metropolis sampler (Metropolis et al. 1953), described in detail elsewhere (Mosegaard and Tarantola 1995; MacKay 2005), except that the sampler's proposal covariance is continuously updated using the covariance of samples in the chain. This process ensures that the proposal covariance (which controls the random walk step size) is optimal for the underlying probability density under consideration. Because of this adaptive characteristic, the sampler's proposal covariance depends on all previous steps, and is not, strictly speaking, a Markov chain. In practice, AM has all other pertinent qualities of the Metropolis MCMC sampler, and so we refer to it, slightly inaccurately, as an MCMC method.

As described in Part I, posterior probability over the parameter vector \mathbf{x} [here the BOSS parameters a , β , and δ from (2)–(5), with a different parameter set for different configurations of BOSS] being sampled is defined by Bayes's theorem,

$$P(\mathbf{x}|\mathbf{y}, \mathcal{M}) = \frac{P(\mathbf{y}|\mathbf{x}, \mathcal{M})P(\mathbf{x}|\mathcal{M})}{P(\mathbf{y}|\mathcal{M})}, \quad (6)$$

where \mathbf{y} is the vector of constraining observations. The term $P(\mathbf{x}|\mathcal{M})$ is the “prior” PDF and represents any a priori quantitative information or belief on the likely values of the parameters. For example, in the current experiment we choose a uniform distribution for the prior, where no parameter value is a priori considered more probable than any other. The term $P(\mathbf{y}|\mathbf{x}, \mathcal{M})$ is the likelihood, and is the term that quantifies comparison between the model and the observations in the space of the observational variables. All quantities in (6) are conditional on the choice of model configuration \mathcal{M} , which in the case of BOSS would be the choice of prognostic moments, or the choice of the number of process rate power-law terms. If expected observational uncertainty is summarized as a multivariate normal distribution, the likelihood is given by

$$P(\mathbf{y}|\mathbf{x}, \mathcal{M}) \sim \mathcal{N}[f(\mathbf{x}, \mathcal{M}), \mathbf{C}], \quad (7)$$

where \mathcal{N} is a multivariate normal distribution with \mathbf{C} as its error covariance and $f(\mathbf{x}, \mathcal{M})$ (the model run with a choice of parameter values \mathbf{x} in configuration \mathcal{M}) its mean value.

The prior for all parameters is a bounded uniform distribution with bounds chosen to enclose a majority of the probability mass, except in cases where further parameter perturbation has little or no influence on BOSS behavior (e.g., if the posterior distribution has constant probability over the range of a parameter). The bounded uniform distribution is defined over the space of $\log_{10}(a)$ and linear β and δ parameters; a parameters are sampled

in log space because a power-law $P = aM^b$ becomes a simple linear function with parameters $\log(a)$ and b upon application of the log transform.

The likelihood defines all sources of uncertainty not present in the prior, and should ideally include

- (i) observational uncertainties (e.g., instrument noise or calibration uncertainties),
- (ii) uncertainties in forward simulation of observable quantities from the model simulation, and
- (iii) structural uncertainties in the model or forward simulator formulations.

Typically only (i) is quantified, and to the extent that (ii) and (iii) remain unquantified, posterior distributions of forward observations $f(\mathbf{x}, \mathcal{M})$ may mismatch constraining observations \mathbf{y} to a degree that is much higher than is quantified in the likelihood uncertainty $\mathcal{N}[f(\mathbf{x}, \mathcal{M}), \mathbf{C}]$. The result of this is a posterior uncertainty that does not accurately describe the actual error in simulation results. We ignore such issues, and tune \mathbf{C} to produce posterior parameter PDFs that are representative of model error. This process is described in the following section.

3. Experimental design

As briefly discussed in the introduction, we present three different experiment types to serve as proof of concept of the probabilistic constraint of BOSS using observations within a Bayesian inference framework. By using idealized models, and a simplified testing framework, we for now ignore the difficulties in using real observations to inform a microphysics scheme [such as model initial condition, boundary condition, and representativeness errors, e.g., as described in [Hodyss and Nichols \(2015\)](#)]. We instead look to highlight challenges in estimation of BOSS parameters in the ideal case where these issues do not exist, or have been adequately addressed. The evaluation of a new approach like BOSS benefits from a modest set of first tests via well-controlled experiments, and we therefore do not seek to address all aspects of the problem of estimating microphysics from observations here.

We expect that a single rainshaft case will be insufficient to constrain BOSS, as not all microphysical processes are likely to exert a dominant and observable influence on evolution of the DSD in a single case ([Kumjian and Prat 2014](#)). To determine the number of cases sufficient to constrain BOSS, we use the criterion that independent choices of column model boundary and thermodynamic conditions should produce equivalent estimates of BOSS parameters. Simply put, we posit that best-estimate BOSS parameter spread from independent

choices of boundary and thermodynamic conditions should not be greater than the spread described by the posterior $P(\mathbf{x}|\mathbf{y}, \mathcal{M})$. We quantify this criterion by comparing the uncertainty within a choice of constraining cases (average standard deviation of the posterior PDFs—“within chain” uncertainty) with the spread between different choices of cases (standard deviation of parameter posterior PDF MAP values and standard deviation of parameter posterior PDF mean values—“interchain” uncertainty). We perform these tests only using M0-M3-B for the sake of simplicity.

[Figure 2](#) shows these statistics for the default number of cases (40) and default choice of observational error \mathbf{C} (corresponding to a standard deviation of 5 dB for moment profiles, 2 dB for column-bottom moment fluxes, top panel). Also, we consider a test of 40 cases with decreased observational error (half of the default standard deviation values—middle panel), and an additional test using 20 cases with the default observational error \mathbf{C} (bottom panel). The statistics using the default number of cases and observational error show the desired behavior, with within-chain parameter deviation (dark bars) greater-than or equal-to interchain deviation of parameter PDF MAP and mean values (light and medium bars, respectively). Reducing the observational error (center panel) results in interchain (MAP and mean) deviation being higher than within-chain deviation, indicating that uncertainty of parameter PDFs does not accurately represent true model parameter certainty, in essence indicating that the posterior PDFs underestimate model error. Reducing the number of cases from 40 to 20 has a relatively mild effect on these statistics, with generally favorable results. However, the 40 cases show lower interchain deviation than 20 cases, in particular for deviation of chain mean. Visual inspection of each parameter PDF (not shown) confirms that for 40 cases, different choices of initial condition samples produce parameter PDFs that are more similar to one another than using 20 cases. Fewer (10) and more (80) cases were tested (not shown); both showed worse characteristics than for either 20 or 40 cases. For these reasons, we use 40 cases for constraining BOSS in results shown hereafter. The method used here is somewhat ad hoc, but is adequate given the idealized nature of the observational constraint. Future work will more thoroughly explore the question of sufficient observational constraint within the context of more realistic observations, including polarimetric radar.

a. Known-truth experiment

We first test BOSS for a case where there is no structural uncertainty—where the constraining observations are themselves generated from a BOSS model

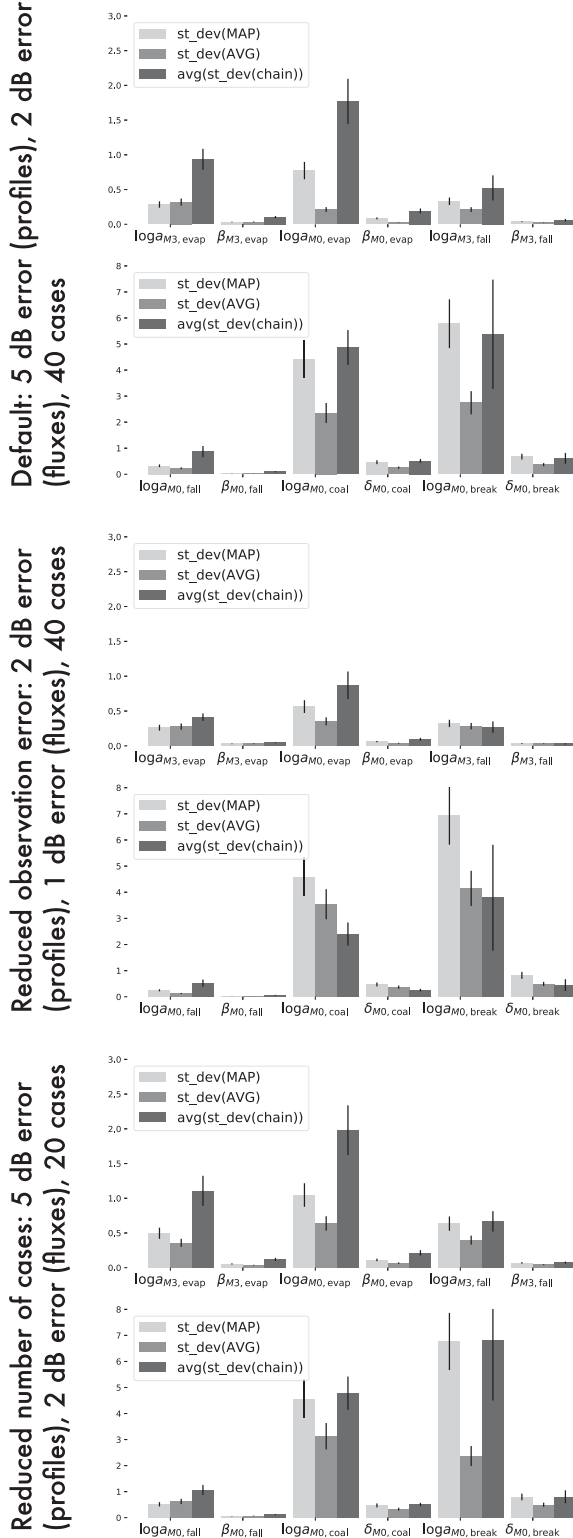


FIG. 2. Statistics of convergence of BOSS parameter PDFs under variation of cases selected for three choices of observational error standard deviation and number of cases used to constrain BOSS. Lighter bars show standard deviation of MAP and parameter PDF average values, respectively, for different

realization. This is deemed a “known-truth” experiment because the true sensitivities of observed microphysical behavior should in principle be reproduced with correct estimation of parameter values. In this case MCMC should accurately find the true parameter values—that is, MAP parameter values should match true parameter values—and uncertainty should only be a product of the nonlinear inverse mapping of observational uncertainty onto the parameter space (see, e.g., Vukicevic and Posselt 2008). For these experiments, observational uncertainty is set arbitrarily low (half of default values) to minimize effects of marginal integration in interpretation of the parameter PDFs. We test the simplest implementation of our model, the one-term two-moment BOSS, with M_0 and M_3 treated as prognostic (M0-M3-B). The truth parameters used to generate the observations are set to MAP values from constraint by MORR described in the experiment below; in other words, they optimally emulate MORR behavior.

b. BOSS model, MORR truth

In the second set of experiments, we consider cases where there is some degree of structural error owing to a mismatch between the “truth” and model microphysics. The performance of configurations of BOSS with various degrees of structural complexity is explored, with all parameters constrained using synthetic observations drawn from rain-shaft simulations using MORR. We compare the performance of three configurations of BOSS: one-term two-moment BOSS with prognostic M_0 and M_3 (M0-M3-B), two-term two-moment BOSS with prognostic M_0 and M_3 (M0-M3-B2), and one-term three-moment BOSS with prognostic M_0 , M_3 , and M_6 (M0-M3-M6-B). In each case, observations of prognostic microphysical variables (M_0 and M_3 for M0-M3-B and M0-M3-B2, M_6 additionally for M0-M3-M6-B) are used at 15 levels in the simulated rain-shaft. The two-term configuration of BOSS (M0-M3-B2) has two terms only for evaporation processes to reduce the number of free parameters, and to reflect the fact that evaporation is described in MORR via two distinct terms, with the second describing ventilation effects on evaporation (see the appendix in Part I).

←

independent choices of case initial conditions. Dark bars show average standard deviation of parameter PDFs for each independent choice of case initial conditions. Shown are statistics (top) for the default values of observational error standard deviation and number of constraining cases, (middle) with reduced observational error, and (bottom) reduced standard deviation.

After the BOSS posterior parameter PDFs have been estimated using MCMC, one thousand parameter samples are chosen from these PDFs, and BOSS and MORR are run again for 10 cases whose initial conditions have been drawn from the climatological database of M_0 , M_3 , M_6 , and RH. This is done so that BOSS can be evaluated for cases that are independent of those that were used to constrain BOSS parameters. BOSS prognostic moments at all levels are output for comparison to MORR, and profiles of microphysical process rates are also output from both BOSS and MORR. As our goal is for BOSS to “learn” correct microphysical process rates from observational constraint, this comparison provides information on the feasibility of such constraint. Finally, these one thousand parameter samples are run for 100 cases to generate more thorough summary statistics of BOSS process rate error relative to MORR.

c. Two-term known truth

As is described in more detail in [section 4](#), testing of M0-M3-B2 revealed that this configuration of BOSS performed no better than M0-M3-B, and often slightly worse (as judged by maximum likelihood values). This occurred despite M0-M3-B2 having more degrees of freedom to fit to data, and despite the fact that the “true” solution for evaporation in MORR is constructed from two distinct power terms. It was found after extensive testing that suboptimal performance in using MCMC to constrain M0-M3-B2 arises from poor MCMC sampler performance caused by strongly non-Gaussian parameter PDFs, with little compelling performance advantage of M0-M3-B2 over M0-M3-B. These problems are in general symptomatic of overfitting. As will be shown, they specifically occur for certain classes of problems involving constraint of parameters in generalized power expressions with multiple terms. To explore these issues in detail, we consider an abstracted form of this problem using the following idealization of the generalized power expressions used for the BOSS process rate formulations:

$$P = \sum_{l=1}^L a_l M_l^b, \quad (8)$$

where L is the number of power-law terms being considered. “Observations” are values of P for 10 values of M evenly spaced between 0.1 and 1 with observational error standard deviation of 0.1. These tests emulate fundamental aspects of constraint of BOSS that arise in the tests described in [section 2b](#).

Ideally, if the constraining observations are generated via (8) with $L = 2$, then estimation should be improved when the model being fit—also using (8)—has $L = 2$. We test this by using MCMC to estimate parameters for

various choices of “true” parameter values (i.e., the ones used to generate observations), and various choices of the model being fit [i.e., varying L in (8)]. We consider three cases where the true model is two term: one where the functional response P is positive definite and monotonic over the values of M considered, another where P is positive definite and nonmonotonic, and a third where the functional response is not positive definite or negative definite, but is monotonic. In each case we consider underfitting (model $L = 1$), perfect fit (model $L = 2$) and overfitting (model $L = 3$). We show posterior parameter PDFs and values of the modeled and true functional response P (see [section 4c](#)).

4. Results

a. Known-truth experiments

Parameter PDFs of the known-truth experiments (M0-M3-B constrained by observations of M0-M3-B) are shown in [Fig. 3](#). In the case of all parameter values, the MAP solution (in blue) is nearly identical to the true parameter values (in red). Further, the peak of the marginal distribution is also nearly identical to the true and MAP parameter values. These results indicate that MCMC is able to unambiguously recover the true parameter values when the structure of the constrained model matches the true model.

Correlated uncertainty between parameters occurs when there is some possibility for compensating errors in parameter values. This is evident in [Fig. 3](#) between matching pairs of $\log(a)$ and β or δ parameters—likely owing to the linear relationship between $\log(a)$ and β or δ parameters in the BOSS process rate equations. Other correlations can be explained by physically plausible mechanisms for compensation of errors. For example, the weak positive correlations between evaporation parameters [most noticeably $\log(a_{M0,\text{evap}})$ and $\beta_{M0,\text{evap}}$] and fall speed parameters [especially $\log(a_{M0,\text{fall}})$ and $\beta_{M0,\text{fall}}$] can be explained by noting that a weaker evaporation rate coupled with a slower fall speed can produce similar vertical profiles of DSD moments as stronger evaporation and faster fall speed. Also noticeable is a negative correlation between coalescence and breakup parameters. These processes have opposing effects, with coalescence decreasing M_0 and breakup increasing M_0 , and thus an underestimation of coalescence could potentially be offset by an underestimation of breakup. However, the correlation between respective a parameters is negative and is, thus, unexpected. It is possible that the correlation is a manifestation of coalescence and breakup being jointly correlated with a third process [analogous to the trivariate relationships explored in [van Lier-Walqui et al. \(2014, Fig. 4\)](#)], and in

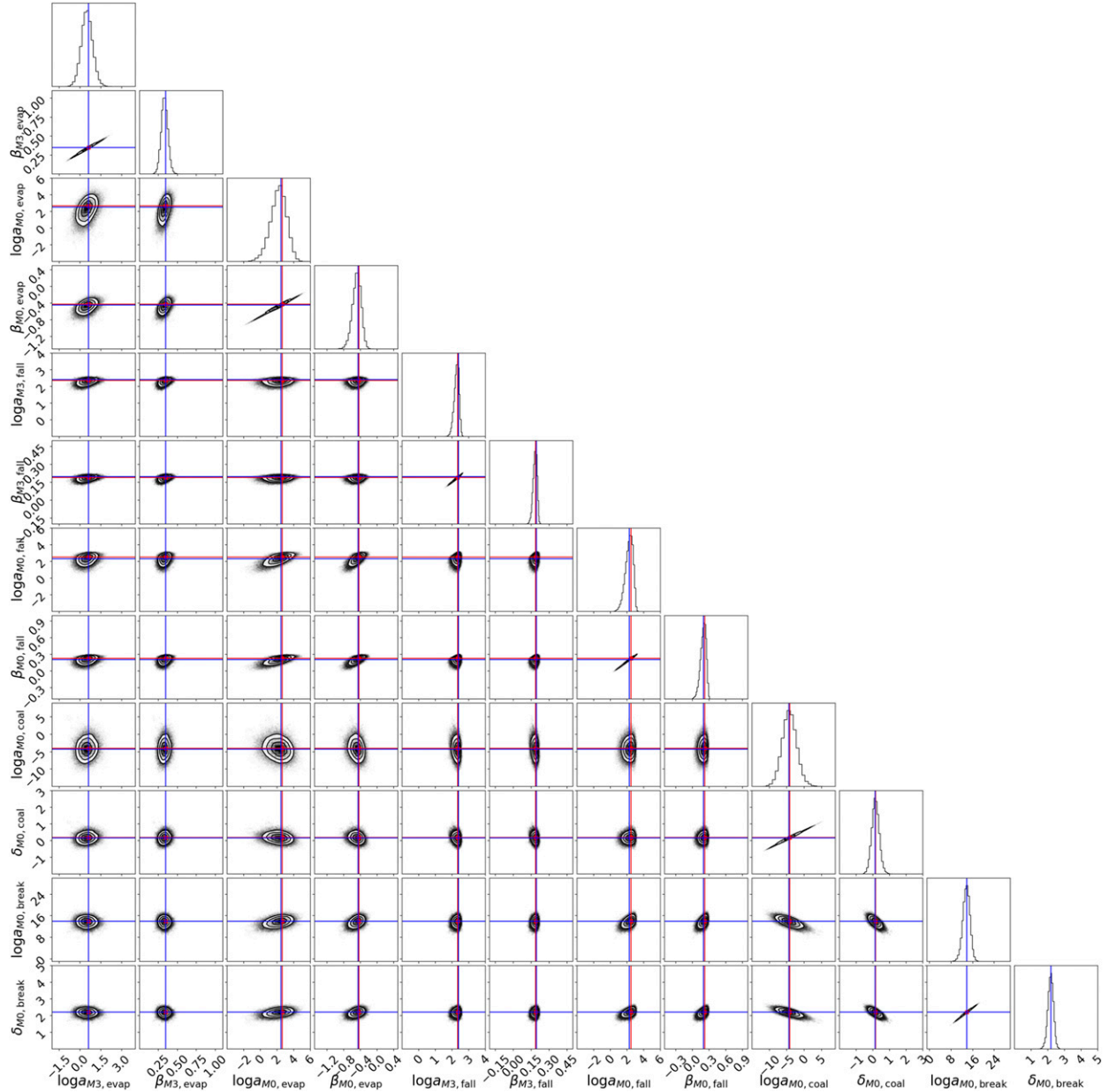


FIG. 3. Posterior parameter PDFs from the known-truth experiment with two-moment (M0, M3), one-term BOSS. “True” parameter values are shown in red, and MAP parameter values from MCMC are shown in blue (where only blue is visible, crosshairs overlap). The parameter PDFs are joint 2D marginals for the parameters as indicated, and 1D marginals along the diagonal.

this context we note that breakup and coalescence parameters show some weak negative and positive correlations with $\log(a_{M0,fall})$ and $\beta_{M0,evap}$. Correlations are also important because they indicate where reduction in uncertainty in one parameter or process is tied to reduction of uncertainty in other parameters.

b. BOSS model, MORR truth

As mentioned previously, results showed virtually identical performance between M0-M3-B and M0-M3-B2,

despite the latter’s increase in structural complexity with two terms in the power expression for evaporation. In many cases, M0-M3-B2 performed notably worse than M0-M3-B from the perspective of MAP log likelihood. For this reason, results are not shown for M0-M3-B2, and a detailed exploration of the reasons for this behavior are discussed in [section 4c](#). Here we focus solely on comparison between M0-M3-B and M0-M3-M6-B.

A basic measure of BOSS performance is shown in [Fig. 4](#): the posterior distribution of $-[f(\mathbf{x}) - \mathbf{y}]^T \mathbf{C}^{-1} [f(\mathbf{x}) - \mathbf{y}]$,

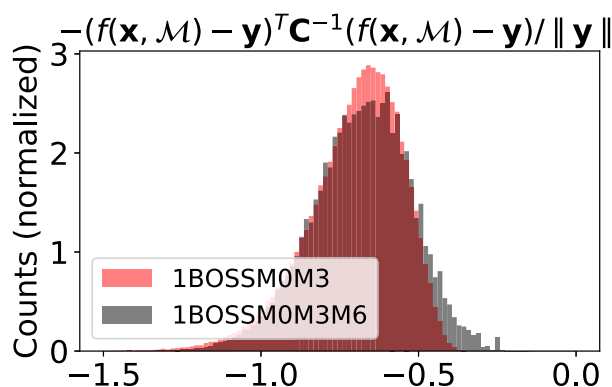


FIG. 4. Posterior distribution of nonnormalized log-likelihood $\{-[f(\mathbf{x}) - \mathbf{y}]^T \mathbf{C}^{-1} [f(\mathbf{x}) - \mathbf{y}]\}$ of the BOSS prognostic moment output for M0-M3-B and M0-M3-M6-B, scaled by the observation-vector degrees of freedom (length of the observation vector \mathbf{y}). Higher values indicate smaller error and hence a better match to observations.

scaled by the degrees of freedom of the likelihood function (i.e., the length of the observation vector). The maximum value of this metric is higher for M0-M3-M6-B (-0.157) than for M0-M3-B (-0.350), indicating better match to observations. The distribution of this metric, while broadly similar between the two BOSS configurations, also displays a greater number of high values for M0-M3-M6-B. Figure 4 thus summarizes the general finding that M0-M3-M6-B outperforms M0-M3-B—an expected result given the exact match in prognostic moments between MORR and M0-M3-M6-B.

The parameter PDFs, shown in Fig. 5, are an MCMC estimate of the probability $P(\mathbf{x}|\mathbf{y}, \mathcal{M})$ from (6) for M0-M3-B (figure for M0-M3-M6-B not shown). Correlations between parameters are generally the same as those found in the known-truth experiment (Fig. 3), indicating that those characteristics of the parameter uncertainty are unchanged when BOSS is constrained by observations that are not produced by BOSS. As in the known-truth case, parameter marginal PDFs are approximately Gaussian, with mild skewness observed in, for example, fall speed parameters such as $\log(a_{M0, \text{fall}})$. One possible explanation for skewness in these parameters is the hard upper limit of 10 m s^{-1} imposed on the highest moment (here M_3) fall speed, as well as the corresponding requirement that lower moments (here M_0) fall more slowly (see Part I). However, no a priori lower-limit constraint exists for fall speed, and thus the constraint on too-low fall speeds must come from observations—we hypothesize that this asymmetry may introduce skewness in the parameter PDFs.

One thousand samples of parameter values are then randomly drawn from the posterior parameter PDFs to run 10 independent cases of 1000 simulations

each—essentially these are a set of “evaluation” cases that are independent of the 40 “training” cases. Profiles of prognostic moment output for these cases are shown in Figs. 6 and 7 for M0-M3-B and M0-M3-M6-B (black lines), respectively, together with rain-rate histograms. These figures also show comparable profiles using MORR (red lines), and profiles using the MAP BOSS parameter values (blue lines). With a few exceptions, results indicate that M0-M3-M6-B has similar prediction of M_0 compared with M0-M3-B. Some notable exceptions such as case 7 show the M0-M3-B ensemble more centered on the MORR profile than the M0-M3-M6-B ensemble, but with a poorer MAP match to MORR. In many instances, such as cases 1, 3, 7, and 9, there is more M_0 uncertainty in M0-M3-M6-B profiles than for M0-M3-B. Overall, Fig. 4 suggests generally better performance for M0-M3-M6-B, but the greater scheme degrees of freedom compared with M0-M3-B may be the cause of increased uncertainty in certain moments for certain cases.

Regarding M_3 profiles, there is some evidence, such as in cases 2, 4, and 6, of M0-M3-M6-B outperforming M0-M3-B, with the case 2 profile a particularly notable example where the entire M0-M3-B ensemble and its MAP value underestimates the MORR M_3 profile. This is also clearly visible in the underestimation of rain rate for case 4 for M0-M3-B; M0-M3-M6-B has a relatively unbiased estimate of rain rate for this case. For M0-M3-M6-B, MAP profiles of M_6 are generally good estimates of MORR M_6 , although in many cases (e.g., cases 1, 3, and 9) the ensemble displays considerable bias in its mean relative to MORR. This bias in ensemble mean indicates that BOSS parameter uncertainty tends to produce consistently smaller M_6 values for these simulations. However, the fact that the MAP profile is a good approximation of MORR indicates that BOSS has the flexibility to reproduce well MORR’s evolution of DSD moment profiles. More generally, multidimensional distributions that are non-Gaussian may have marginal distributions with modes that differ from those of the PDF in its full “native” dimensions—see Vukicevic and Posselt (2008) for illustrative examples of this phenomenon.

Figures 8 and 9 are generated similarly to Figs. 6 and 7, and show profiles of microphysical process rates from the 10 validation cases using M0-M3-B and M0-M3-M6-B, respectively. All microphysical process rates are output at every level in the model column from both BOSS and MORR. The processes of collision-coalescence and collisional breakup are combined (coalescence-breakup) to match how these processes are treated within MORR. However, negative and positive values of coalescence-breakup are shown separately to distinguish between breakup- and coalescence-dominated cases, respectively. This also allows for displaying the rates in dB, which more

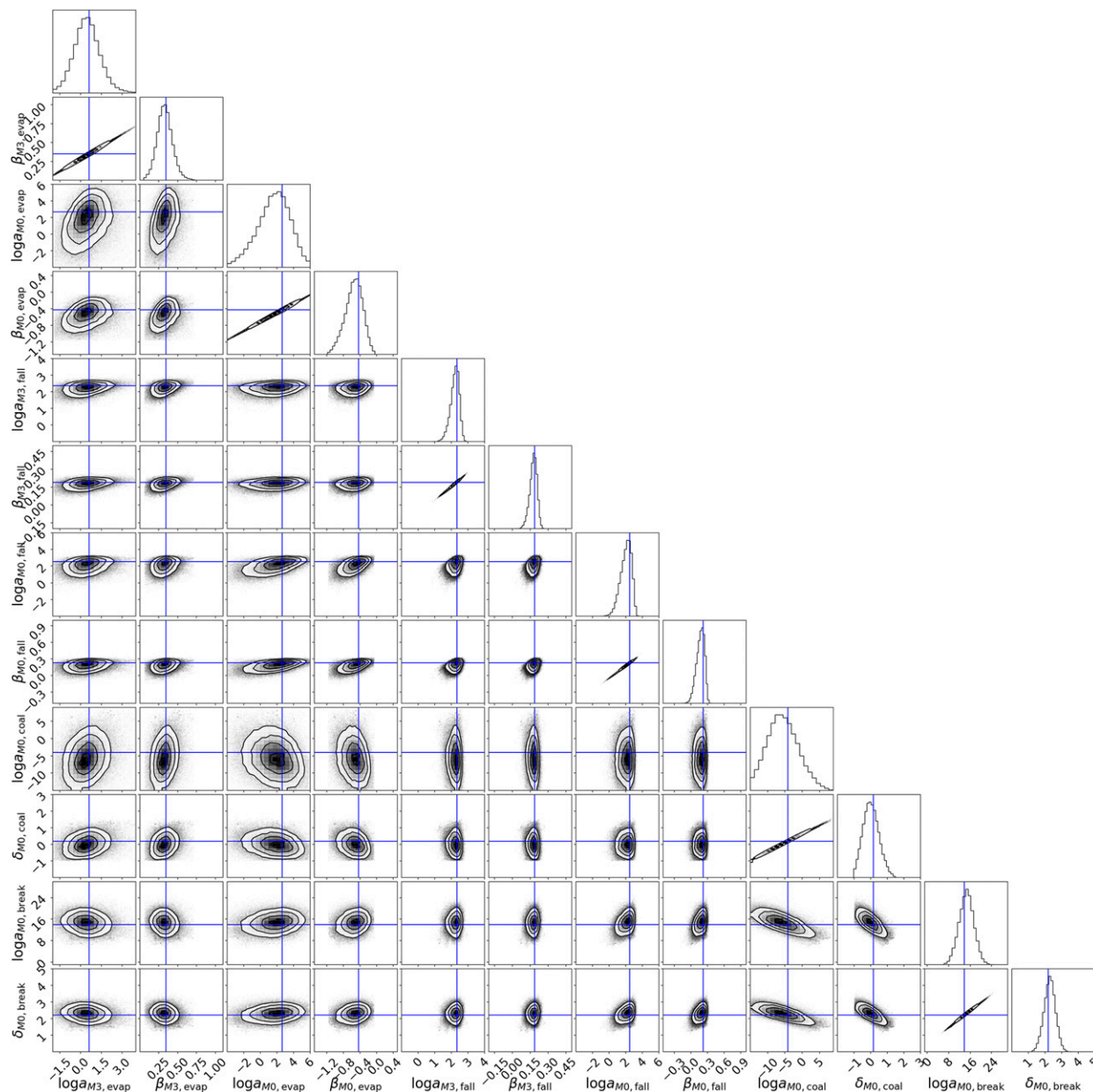


FIG. 5. Posterior parameter PDFs for M0-M3-B, constrained by MORR. MAP parameter values are indicated by blue crosshairs. The parameter PDFs are joint 2D marginals for the parameters as indicated, and 1D marginals along the diagonal.

adequately shows the full range of values simulated (which can span multiple orders of magnitude between column top and column bottom). Values of coalescence and breakup rates for M_6 are multiplied by -1 so that positive values of coalescence–breakup for M_6 match with positive values of coalescence–breakup for M_0 . Moment fall speeds are presented in meters per second.

In effect, Figs. 8 and 9 demonstrate the degree to which BOSS has “learned” microphysical process rates via Bayesian constraint from bulk rain observations.

We note that for both M0-M3-B and M0-M3-M6-B, BOSS process rates are generally constrained to within 1–2 dB of MORR. Furthermore, the shape of the MORR process rates profiles are well captured by BOSS process rates—where process rates increase/decrease with height in MORR, they tend to do so in BOSS as well. The transition between breakup- and coalescence-dominated regions of the profile, shown in the balance between profiles with positive and negative coalescence–breakup, is also well represented in BOSS when it exists

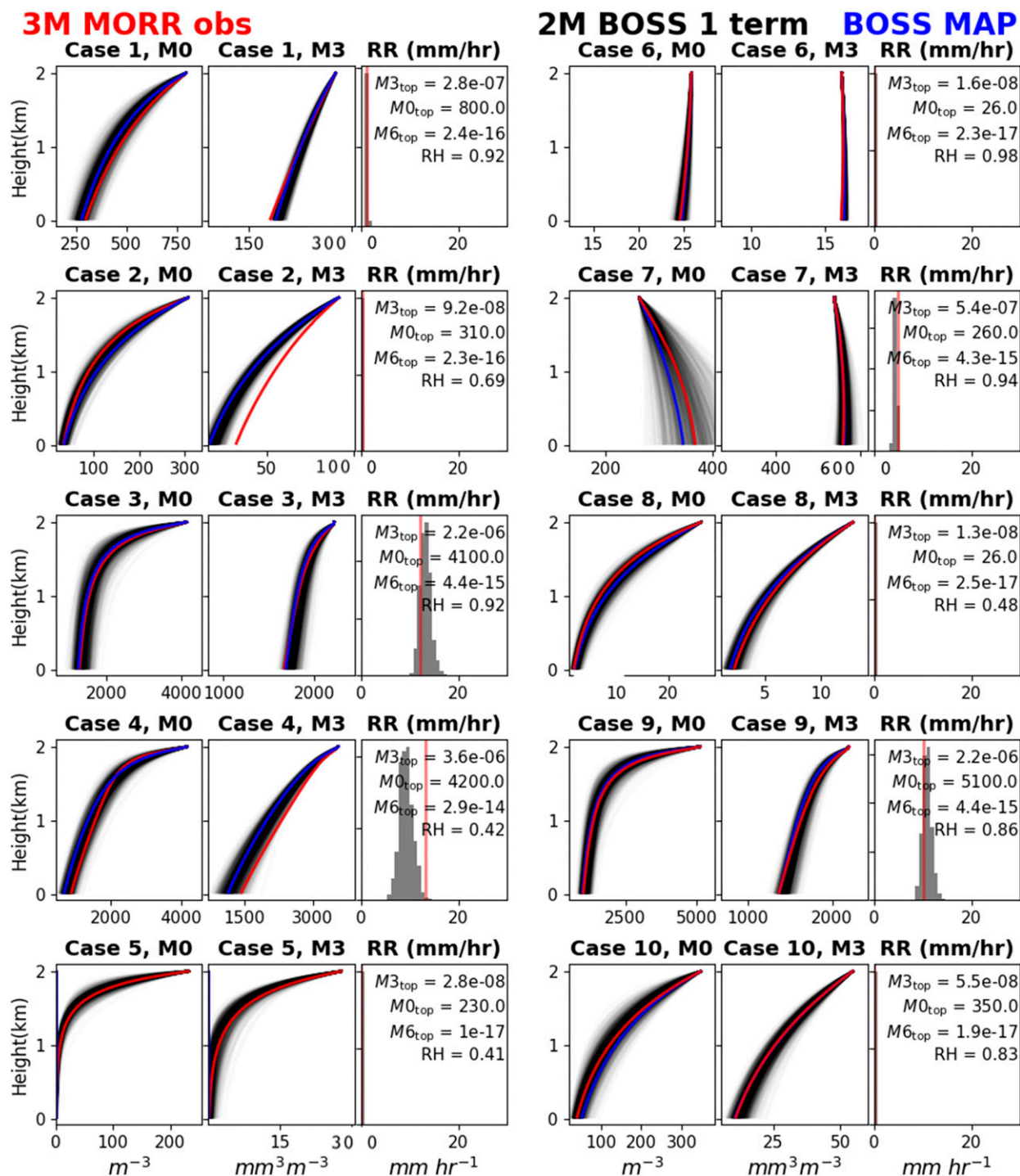


FIG. 6. Profiles of prognostic moments (M_0 , M_3) and rain-rate histograms, for 10 cases where the model is run using BOSS with 1000 parameter values sampled from the M0-M3-B PDF (black), using BOSS with the M0-M3-B MAP parameter values (blue), and using MORR (red). Specified upper boundary conditions sampled for each case are shown by text in the rain-rate histogram plots.

for MORR (e.g., cases 3, 4, and 9), with some uncertainty in the height at which this transition occurs. For example, for M0-M3-B case 4 this transition occurs about 400 m too low in the BOSS MAP profile, but is much better

estimated in case 3. We note that in cases where process rates approach 0, strongly underestimated values may still be insignificant, because the scale shown is in log units (i.e., a 100-fold underestimation of an insignificantly small

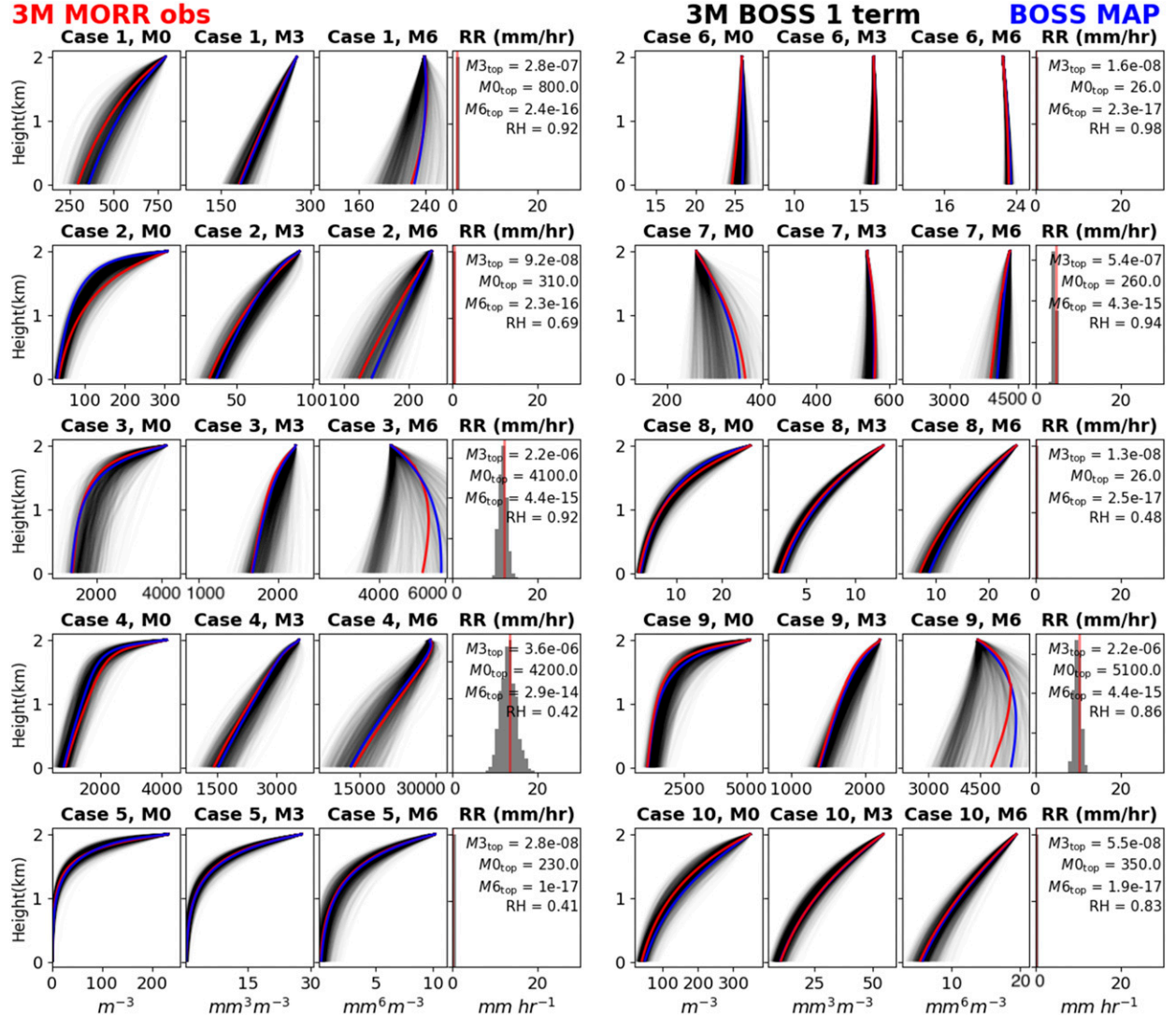


FIG. 7. As in Fig. 6, but for M_0 , M_3 , and M_6 using the three-moment M0-M3-M6-B. The y axes are identical to Fig. 6, but the x-axis ranges vary.

number is still an insignificantly small number). This, in part, explains the increase in BOSS spread for large negative dB values of $M0_{coal}$ shown in Fig. 8 for cases 3, 4, and 9.

Comparison of the results from M0-M3-B, shown in Fig. 8, and from M0-M3-M6-B, shown in Fig. 9, displays tradeoffs in how each configuration of BOSS gains process-level information from the observational constraint. For example, while M0-M3-M6-B better estimates MORR profiles of V_{M3} , most readily visible in cases 1, 4, 5, 6, and 10, M0-M3-B better characterizes $M0_{coal}$ and $M0_{break}$, especially for cases in which the rainshaft transitions from coalescence-dominated evolution aloft to breakup dominated near the surface (cases 3, 4, and 9).

Figure 10 shows summary histograms of BOSS process rate error (dB) for all levels and all simulations. Here to produce more robust statistics we run 100 cases initialized from the same database as before. In virtually all cases, the PDFs of error are nearly symmetric about 0, indicating that estimation of process rate is, in the mean, mostly unbiased. Furthermore, with the exception of $M6_{coal}-M6_{break}$ and $M0_{coal}-M0_{break}$, process rate error is mostly less than 2 dB for all processes. Comparing results from M0-M3-B and M0-M3-M6-B, we find virtually identical results for $M3_{evap}$, a slight advantage for M0-M3-B in $M0_{evap}$, V_{M0} , and V_{M3} , and no clearly better choice for $M0_{coal}-M0_{break}$; these process error distributions have generally equal probability of 0-dB error, but differ in mean error and error standard

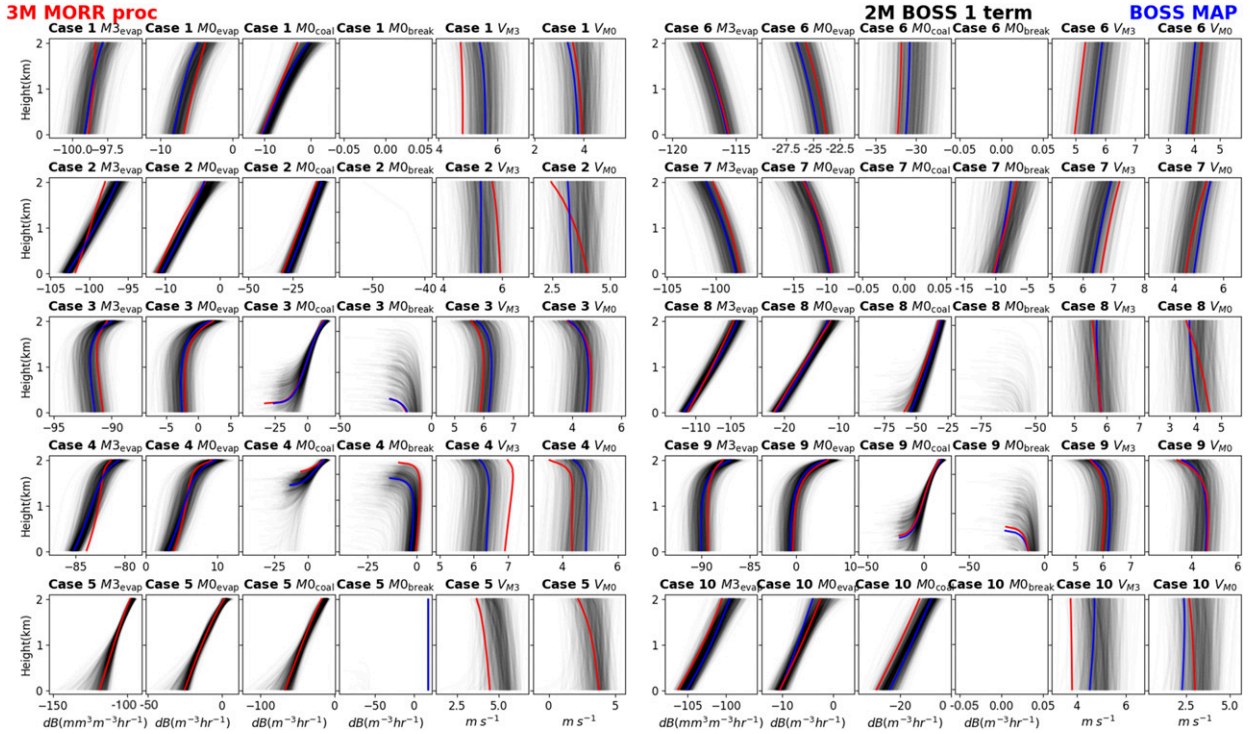


FIG. 8. As in Fig. 6, but for process rate output from each respective simulation.

deviation between M0-M3-B and M0-M3-M6-B. We note that this error distribution is effectively the sum of two errors (from coalescence and breakup separately) and thus its variance is the sum of each individual variance.

c. Two-term known truth

As discussed previously, tests using M0-M3-B2 generally showed no improvement over M0-M3-B, despite the increased degrees of freedom with which to fit BOSS behavior toward MORR. Furthermore, owing to strong non-Gaussianity of the M0-M3-B2 parameter PDF, the performance of the MCMC sampler was considerably reduced, which often resulted in *worse* performance of M0-M3-B2 relative to M0-M3-B. This was true even when M0-M3-B2 was used to generate the constraining observations—that is to say, even when the truth was generated using a two-term power law for evaporation, a one-term model often provided a better fit (results not shown here for brevity). Upon further investigation, this issue was found to be emblematic of a certain class of problems where parameters of generalized power expressions are estimated. This is illustrated by examining the model response P as a function of independent variable M following the idealization of the BOSS process rate power expressions in (8). We show one case that is representative of the problems encountered

in estimating M0-M3-B2 parameters, as well as two contrasting cases where unambiguous estimation of parameters for a two-term power expression is straightforward.

Results for the three cases are shown in Figs. 11–13. The first, Fig. 11, shows the case where the true solution is monotonic and positive definite. The top plot shows the true solution (green) as well as the contribution of each term in the two-term power law (orange and blue). The three rows below show the results of Bayesian parameter estimation using one-term (underfit), two-term (perfect fit), and three-term (overfit) models. For this case, the one-term model shows an approximately Gaussian bivariate PDF for the parameters, and a very good match between the solution of this model and truth. The two- and three-term models also produce good matches to the true solution. However, their respective parameter PDFs are strongly non-Gaussian. In this case, a one-term solution provides a very good approximation of two-term truth while maintaining a well-behaved (e.g., Gaussian) posterior distribution. Strong non-Gaussianity observed for the two- and three-term solutions could result in difficulty in estimating those parameters. This case is similar to that found exploring two terms in BOSS—a one-term solution performed adequately and two-term parameter PDFs were found to be strongly non-Gaussian.

3M BOSS 1 term

3M MORR proc

BOSS MAP

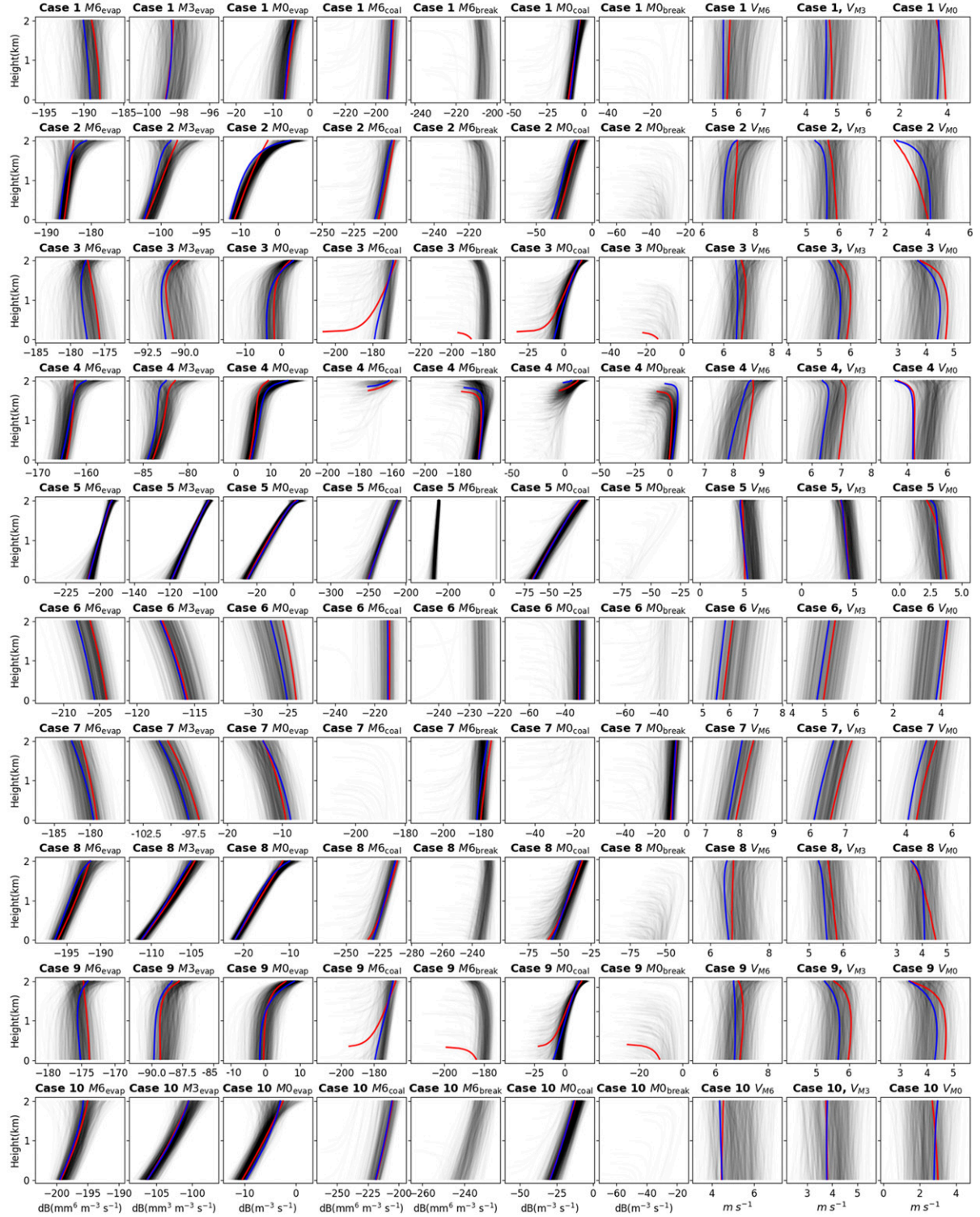


FIG. 9. As in Fig. 8, but for process rate profile output from each respective simulation. The y axes are identical to Fig. 8, but the x-axis ranges vary.

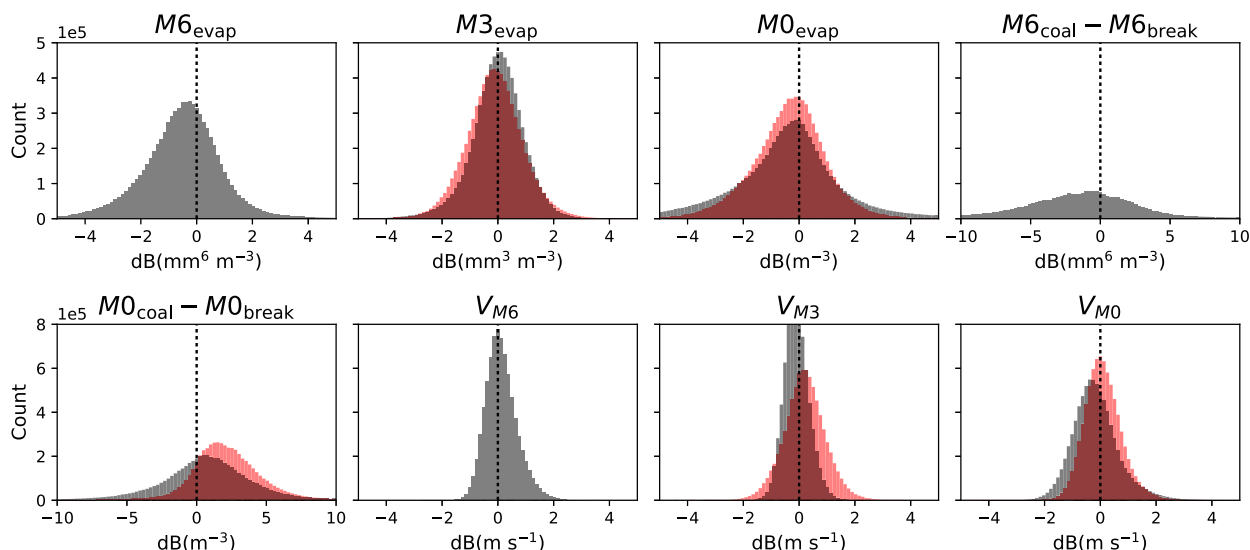


FIG. 10. Process rate error (dB) of BOSS simulations relative to MORR, aggregated for all levels in the column for 100 simulations, which include the 10 simulations shown in Figs. 8 and 9, and 90 additional, independently generated simulations. M0-M3-B and M0-M3-M6-B process rate error histograms are shown in red and gray, respectively. The y axes are consistent across rows.

Truth in the second case (Fig. 12) is marked by a nonmonotonic relationship between the response P and changes in the independent variable M . Again, for this case, the one-term model's parameter PDF is nearly Gaussian, but its solution is a poor approximation of truth—a single power law inherently cannot reproduce this nonmonotonic function. The parameter PDF of the two-term model now displays only mild nonlinearities, with nearly Gaussian one-dimensional marginal distributions; this two-term model produces solutions that closely match truth. The three-term model again displays strongly non-Gaussian parameter PDFs, but its solutions closely approximate truth. In this case, where the two-term model was necessary to approximate the two-term truth, the posterior PDF was closer to Gaussianity than in the first case.

Truth in the third case (Fig. 13) is a monotonic two-term function, but not positive definite (nor negative definite, i.e., it changes sign with M). Again, the one-term model has a Gaussian parameter PDF and solutions that poorly approximate truth, as it inherently cannot reproduce solutions that change sign. The parameter PDF of the two-term model is again moderately non-Gaussian, but generally well behaved (e.g., unimodal), and its solutions are a good approximation of truth. The three-term model again has non-Gaussian parameter PDFs, and solutions that well approximate truth. We note, but do not show here, that this case can also be well described by estimating a_1 , b_1 , and a_2 with b_2 set to 0, thus reducing the number of free parameters from four to three. To summarize, for this third class of

two-term power laws, a model with at least two terms (and at least three or four free parameters) is necessary to reasonably approximate truth.

The first case (Fig. 11, monotonic and positive definite) is quite similar to most microphysical process rates, which we typically would expect to vary monotonically with moment value. For example, an increase in the bulk droplet concentration while maintaining the same DSD shape will increase the value of all moments (i.e., all moment orders), and all else equal bulk process rates will be larger because they are operating on more drops. We therefore expect that for many of these processes (though not all, particularly for higher-order moments), a single power-law term may be sufficient to capture natural behavior. As noted in Part I, the final case shown here (monotonic, not positive definite or negative definite) can be thought of as analogous to the combined processes of collision-coalescence and breakup within BOSS; each process is governed by a separate power law, but has opposite effects on the evolution of DSD moments. Effectively, the combined coalescence-breakup process is not positive definite or negative definite, and thus must be represented by at least two terms, because it is negative in coalescence-dominated regimes and positive in breakup-dominated regimes for the change in M_k when $k < 3$ (vice versa for $k > 3$).

5. Discussion and conclusions

A novel rain microphysics scheme, the Bayesian Observationally Constrained Statistical-Physical Scheme

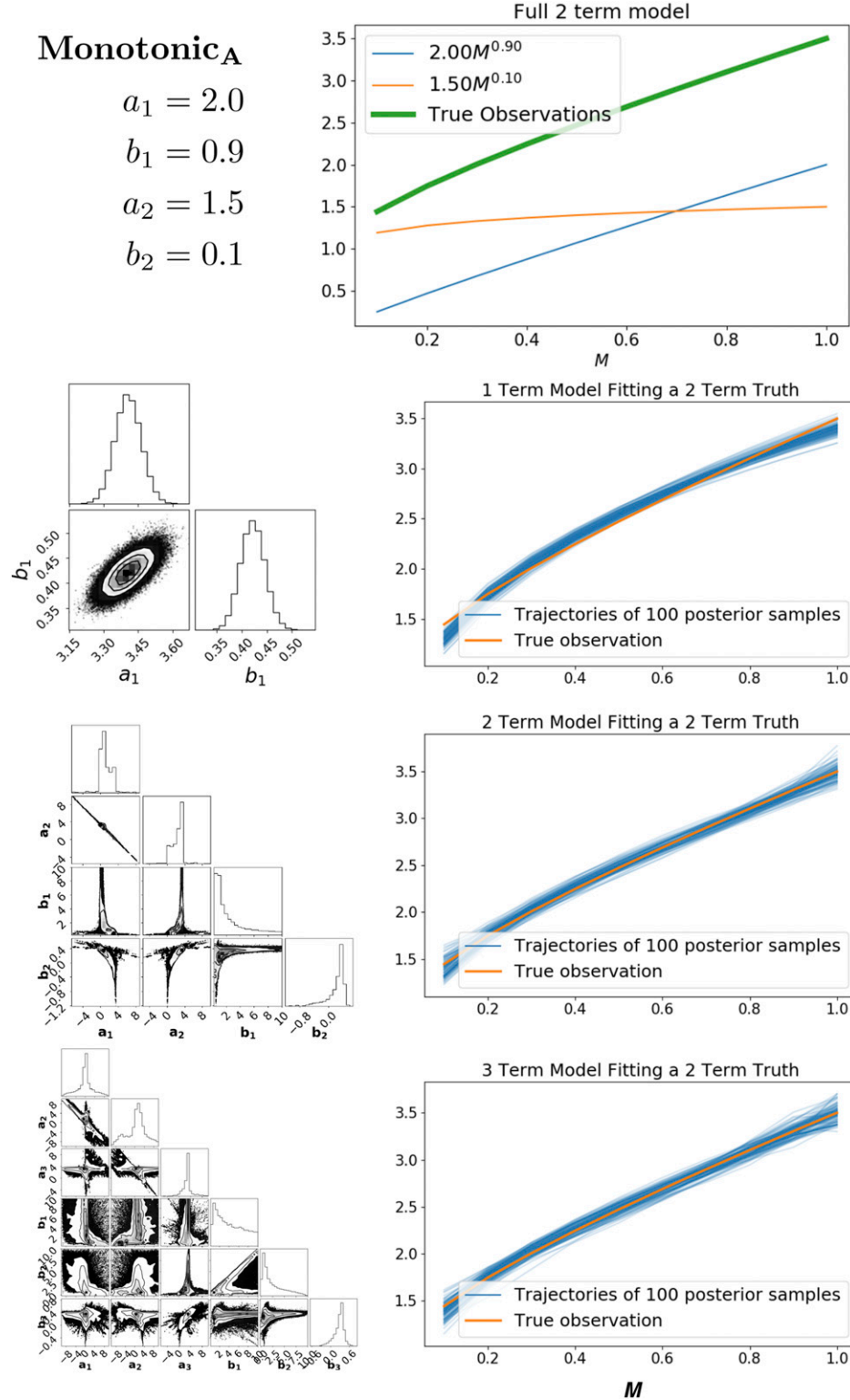


FIG. 11. Tests of estimation of parameters of two- and one-term power-law models for the case in which the true power law is a monotonically increasing two-term power law. (top) In the full two-term “truth,” the blue and red lines representing each power-law term comprising the full solution shown by the green line. Values of the “truth” parameters are to the left of this plot. (bottom-right column) Results using the one-term, two-term, and three-term models for 100 forward calculations using parameters sampled from the posterior PDFs (blue lines) and the “truth” (red lines). (bottom-left column) Posterior parameter PDFs for the one-term, two-term, and three-term models.

Non monotonic

$$a_1 = 5.0$$

$$b_1 = 0.75$$

$$a_2 = 1.5$$

$$b_2 = -0.75$$

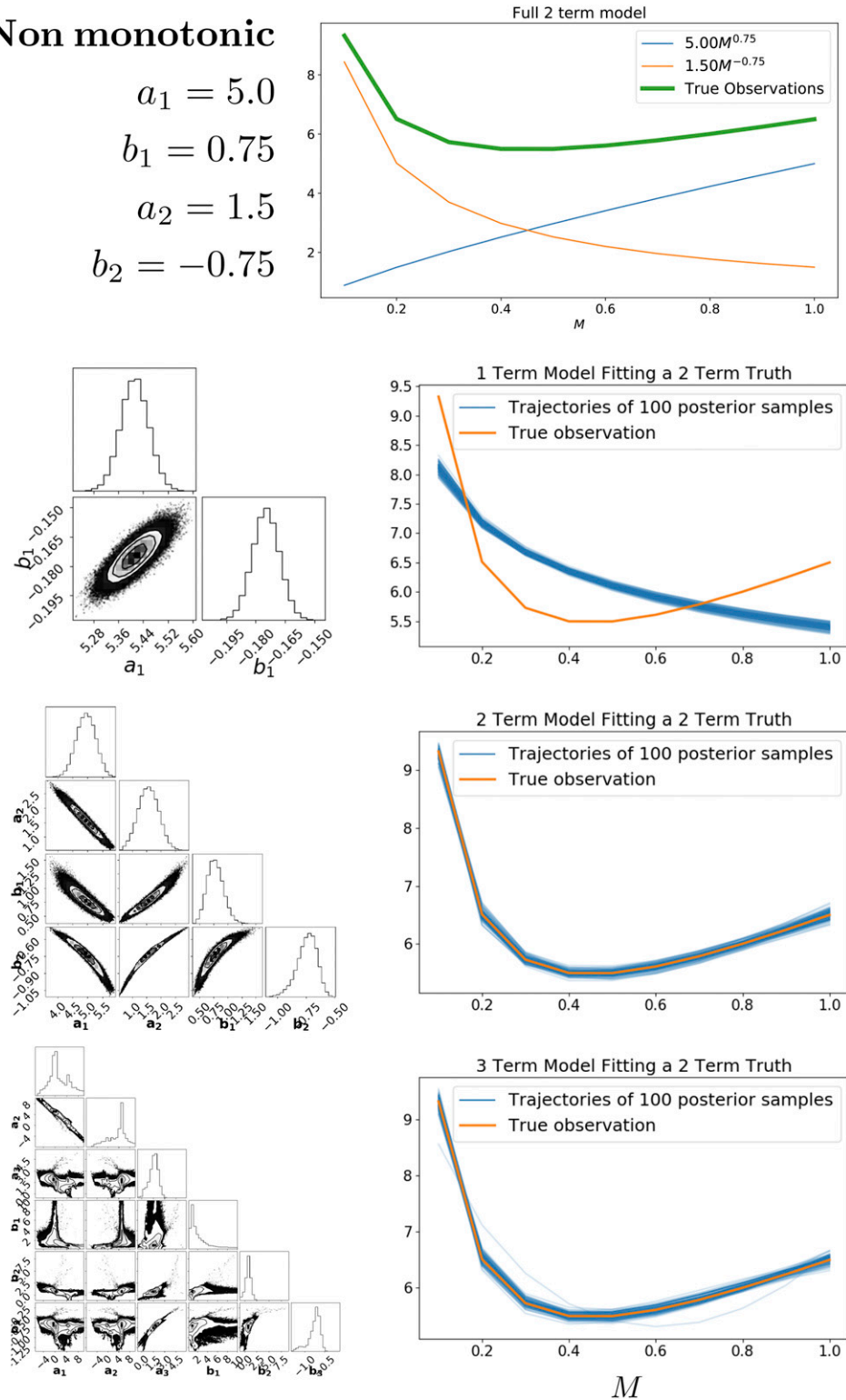


FIG. 12. As in Fig. 11, but for the case that the true power law is a nonmonotonic two-term power law.

Monotonic_B

$$a_1 = 5.0$$

$$b_1 = 0.75$$

$$a_2 = -1.5$$

$$b_2 = -0.5$$

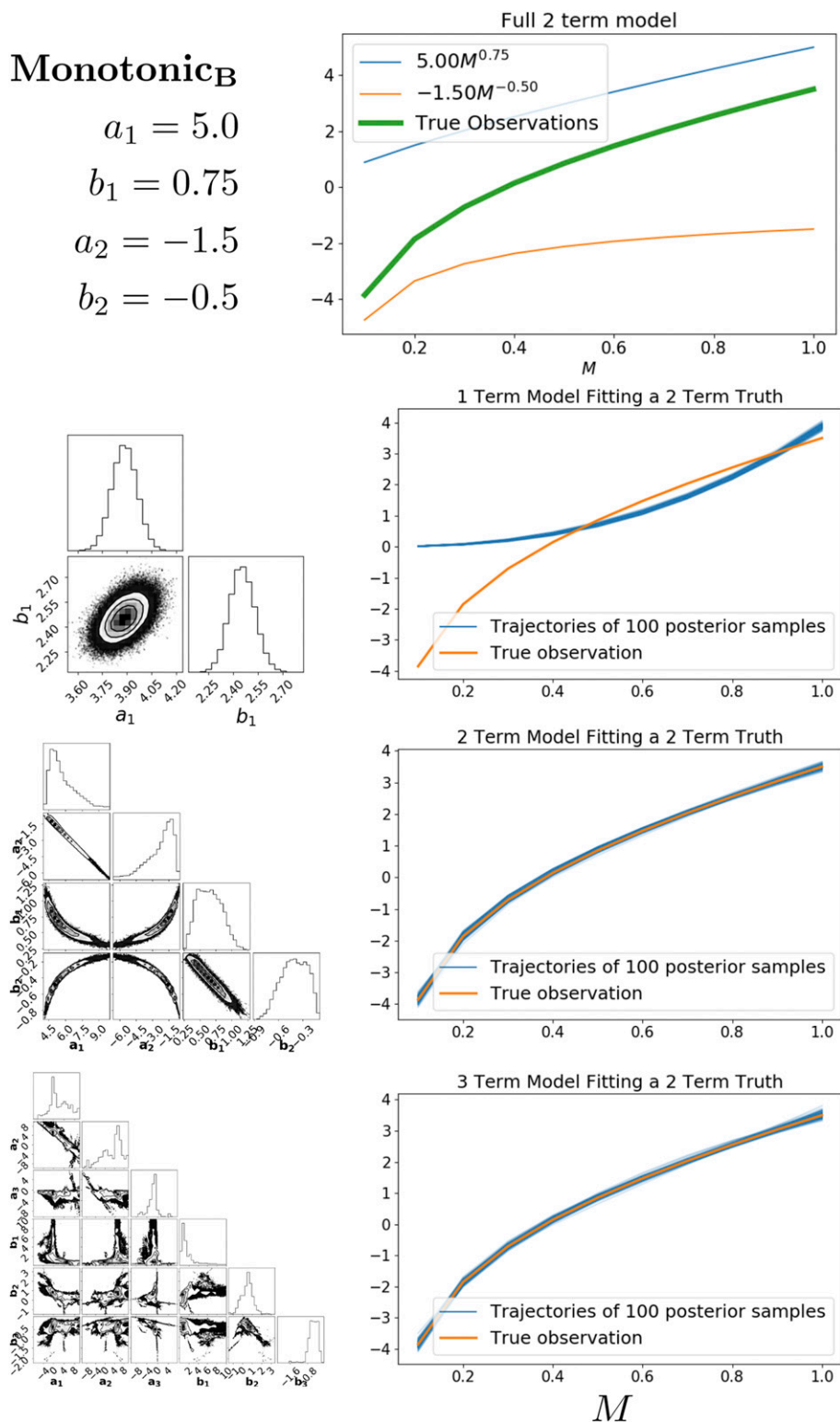


FIG. 13. As in Figs. 11 and 12, but for the case that the true power law is a monotonic but not positive-definite two-term power law.

(BOSS; introduced in [Part I](#)), was constrained within a one-dimensional idealized rain-shaft model using synthetic observations that were generated by a “traditional” three-moment bulk microphysics scheme. Different configurations of BOSS with varying degrees and types of structural complexity were compared to ascertain the possible challenges and potential of developing a microphysics scheme informed almost entirely by observations. All constraint was performed using an MCMC-like Bayesian parameter estimation algorithm that also produced robust estimates of parameter uncertainty.

Results showed that, in general, observations were capable of informing BOSS parameters without significant nonuniqueness in the solutions. Correlated uncertainty between BOSS parameters indicated the possibility for compensating errors in process rates, but generally these effects were weak compared with correlations between $\log(a)$ and β and δ parameters. Correlation can suggest additional observations that could be particularly useful for constraining BOSS. For example, correlation between uncertainty in evaporation and sedimentation parameters indicates that estimation of *both* sedimentation and evaporation parameters would benefit by better constraint of the sedimentation rate, for example, via direct measurements of Doppler fall speed by vertically pointing radars.

It was also noted that collision-coalescence and collisional breakup parameters were somewhat worse constrained than sedimentation and evaporation parameters. We noted that the error variance is effectively the sum of the individual variances of coalescence and breakup. However, this only partly explains the greater uncertainty in coalescence/breakup versus evaporation. An explanation for this increased uncertainty is that, in the case of M0-M3-B, observations of M_3 inform only two processes, evaporation of M_3 and sedimentation of M_3 , whereas observations of M_0 informed four: evaporation, sedimentation, collision-coalescence, and breakup. Collision-coalescence and breakup were therefore more poorly constrained than, say, evaporation of M_3 . This underscores the need for strong observational information relating to the evolution of moments of interest, in particular those that are evolved by *more* parameterized physical processes. These observational needs suggest that M_6 might provide a better companion to M_3 in a two-moment scheme, as radar reflectivity and reflectivity-weighted fall speed are routinely observed by scanning and profiling radar systems, respectively. Another reason for the poorer constraint of collision-coalescence and breakup parameters is the fact that most cases used for observationally constraining BOSS were more strongly controlled by evaporation—[Fig. 8](#) showed that the combined rates of coalescence/breakup

were much lower than that of evaporation for at least six of the cases considered (cf. $M0_{\text{evap}}$ and $M0_{\text{coal}}$ for cases 1, 2, 5, 6, 8, and 10). As such, model evolution was relatively less sensitive to perturbations in coalescence/breakup parameters, resulting in weaker constraint. Hypothetically, this uncertainty could be reduced by observationally constraining using more cases where coalescence and/or breakup rates are at least as large as evaporation rates. Future work will explore the diversity of realistic observations needed to inform each process, something beyond the scope of the current research.

One of the key features of BOSS relative to other schemes is its flexibility in structural complexity; both the number of prognostic moments, choice of prognostic moments, and the number of terms in the generalized power expressions for the process rates can be systematically altered to match observed microphysical complexity. Here, we considered two increases in complexity beyond that of the two-moment BOSS with single-term power-law expressions for the process rates (M0-M3-B): the addition of another prognostic moment (M_6 , in M0-M3-M6-B), and separately, the addition of another term in process rate expressions (M0-M3-B2). The addition of M_6 as a prognostic moment improved overall scheme fidelity, in particular improving prediction of M_3 and rain rate. However, estimation of microphysical process rates (for M_0 and M_3) was not significantly improved compared with M0-M3-B. This was likely a consequence of additional degrees of freedom without a commensurate increase in observational information; the number of free parameters more than doubled from 12 to 30, whereas the observation vector length increased by only 50%. In tests with an appropriately increased number of observations for M0-M3-M6-B (not shown), process errors for M0-M3-M6-B were, for all processes, found to be lower than for both M0-M3-B and M0-M3-M6-B with the default number of observations.

Results from tests adding structural complexity to BOSS by using two power-law terms to model evaporation had little to no positive impact on results, with some MAP estimates noticeably worse than using one power-law term, despite the fact that the true process rate in MORR was a two-term function of the prognostic moments. These results can be interpreted within the context of idealized tests using simplified power laws ([section 4](#)). For certain two-term power laws, which are monotonic and positive definite, a single power-law term was sufficient to adequately model the true function response, and the addition of further terms showed symptoms of overfitting (e.g., strongly non-Gaussian parameter PDFs). Other two-term power-law functions, namely, those that are either nonmonotonic or monotonic but not positive definite or negative definite,

required a two-term model for adequate fit, and only showed symptoms of overfitting for three-term models. These three classes of two-term power laws are neither exhaustive nor necessarily stereotypes for the true functional form of microphysical process rates. Instead they give insight into where problems can arise with fitting power expression models, and yield some limited heuristics to guide interpretation of BOSS parameter estimation results. First, if a two-term model does not show obvious improvement over a one-term model, and the two-term model parameter PDFs are strongly non-Gaussian, then it is likely that the two-term model is overfitting to the truth *even if truth is multiterm*. If the true functional response is nonmonotonic or non-positive (or nonnegative) definite, it is highly likely that at least two terms will be needed to simulate the truth, as is the case for the combined coalescence–breakup process. We leave open the possibility that there exist monotonic, positive- or negative-definite functions that require multiple power laws for adequate modeling of microphysical processes.

Future progress and development of BOSS will ultimately depend on the ability to formulate appropriate parameter inference problems. This involves three interlocking components: the observations used to constrain the model, the model that simulates the observed weather (of which BOSS is one component), and the Bayesian parameter estimation framework. There are crucial considerations over what sort of real observations may be necessary to provide adequate constraint on BOSS process representation. Additionally, the ability to retrieve microphysical information from observations will depend closely on the choice of model used to simulate weather. Here we used a steady-state rain-shaft model, and while there may be some weather phenomena amenable to such treatment (broad, spatially and temporally homogeneous stratiform regions of mesoscale convective systems, etc.), most convective rain will be poorly simulated by a steady-state model. Furthermore, observations from these convective systems may be necessary to constrain microphysical processes such as collision–coalescence and collisional breakup, which tend to be more active at higher rain rates. On the other hand, freely evolving (i.e., nonsteady state) simulations will suffer from error growth associated with both physical uncertainties as well as initial condition uncertainty, with strong coupling between the freely evolving microphysical, dynamical, and thermodynamical states. As such, these cases will likely require simultaneous state and parameter estimation in order to prevent state errors from rendering microphysical errors uninterpretable. No widely used system exists for such simultaneous parameter and state estimation, though candidates

have been proposed (e.g., Laine et al. 2012). Finally, we note that observations should be capable of informing both the most probable parameter values, as well as choice of most probable *model* (or BOSS configuration); that is, they should be capable of informing not only parameter estimation experiments, but also *model selection* experiments. Bayesian techniques for model selection are commonly applied to hydrological (Höge et al. 2018) and seismological (Sambridge et al. 2013) modeling, but have not yet gained traction within the atmospheric sciences.

Although these issues will require concerted efforts to resolve, BOSS is uniquely suited to providing the microphysical modeling framework for these tasks, owing to the flexibility of its structure while retaining a physical framework. Thus, observationally driven development of BOSS also holds promise to yield meaningful physical insights into microphysics on the process level because BOSS retains a physically interpretable structure. In these respects, BOSS has advantages over both traditional microphysics schemes, as well as “black box” approaches to data-driven parameterization development.

Acknowledgments. This work was funded by U.S. DOE Atmospheric System Research Grant DE-SC0016579. The National Center for Atmospheric Research is sponsored by the National Science Foundation. MM and SL acknowledge support by the National Science Foundation under Grant DMS-1619630. MM gratefully acknowledges support by the Office of Naval Research (Grant N00173-17-2-C003).

REFERENCES

- Berner, J., S.-Y. Ha, J. P. Hacker, A. Fournier, and C. Sn, 2011: Uncertainty in a mesoscale ensemble prediction system: Stochastic versus multiphysics representations. *Mon. Wea. Rev.*, **139**, 1972–1995, <https://doi.org/10.1175/2010MWR3595.1>.
- , K. R. Fossell, S.-Y. Ha, J. P. Hacker, and C. Snyder, 2015: Increasing the skill of probabilistic forecasts: Understanding performance improvements from model-error representations. *Mon. Wea. Rev.*, **143**, 1295–1320, <https://doi.org/10.1175/MWR-D-14-00091.1>.
- Diamond, H. J., and Coauthors, 2013: U.S. Climate Reference Network after one decade of operations: Status and assessment. *Bull. Amer. Meteor. Soc.*, **94**, 485–498, <https://doi.org/10.1175/BAMS-D-12-00170.1>.
- Haario, H., E. Saksman, and J. Tamminen, 2001: An adaptive Metropolis algorithm. *Bernoulli*, **7**, 223–242, <https://doi.org/10.2307/3318737>.
- Hodyss, D., and N. Nichols, 2015: The error of representation: Basic understanding. *Tellus*, **67A**, 24822, <https://doi.org/10.3402/tellusa.v67.24822>.
- Höge, M., T. Wöhling, and W. Nowak, 2018: A primer for model selection: The decisive role of model complexity. *Water Resour. Res.*, **54**, 1688–1715, <https://doi.org/10.1002/2017WR021902>.

- Jankov, I., and Coauthors, 2017: A performance comparison between multiphysics and stochastic approaches within a North American RAP ensemble. *Mon. Wea. Rev.*, **145**, 1161–1179, <https://doi.org/10.1175/MWR-D-16-0160.1>.
- Kumjian, M. R., and O. P. Prat, 2014: The impact of raindrop collisional processes on the polarimetric radar variables. *J. Atmos. Sci.*, **71**, 3052–3067, <https://doi.org/10.1175/JAS-D-13-0357.1>.
- , C. P. Martinkus, O. P. Prat, S. Collis, M. van Lier-Walqui, and H. C. Morrison, 2019: A moment-based polarimetric radar forward operator for rain microphysics. *J. Appl. Meteor. Climatol.*, **58**, 113–130, <https://doi.org/10.1175/JAMC-D-18-0121.1>.
- Laine, M., A. Solonen, H. Haario, and H. Järvinen, 2012: Ensemble prediction and parameter estimation system: The method. *Quart. J. Roy. Meteor. Soc.*, **138**, 289–297, <https://doi.org/10.1002/QJ.922>.
- Leeper, R. D., J. Rennie, and M. A. Palecki, 2015: Observational perspectives from U.S. Climate Reference Network (USCRN) and Cooperative Observer Program (COOP) network: Temperature and precipitation comparison. *J. Atmos. Oceanic Technol.*, **32**, 703–721, <https://doi.org/10.1175/JTECH-D-14-00172.1>.
- MacKay, D. J. C., 2005: *Information Theory, Inference, and Learning Algorithms*. 7th ed. Cambridge University Press, 628 pp.
- Metropolis, N., and Coauthors, 1953: Equation of state calculations by fast computing machines. *J. Chem. Phys.*, **21**, 1087–1092, <https://doi.org/10.1063/1.1699114>.
- Moiseev, D. N., S. Lautaportti, J. Tyynela, and S. Lim, 2015: Dual-polarization radar signatures in snowstorms: Role of snowflake aggregation. *J. Geophys. Res. Atmos.*, **120**, 12 644–12 655, <https://doi.org/10.1002/2015JD023884>.
- , A. von Lerber, and J. Tiira, 2017: Quantifying the effect of riming on snowfall using ground-based observations. *J. Geophys. Res. Atmos.*, **122**, 4019–4037, <https://doi.org/10.1002/2016JD026272>.
- Morrison, H., and W. W. Grabowski, 2007: Comparison of bulk and bin warm-rain microphysics models using a kinematic framework. *J. Atmos. Sci.*, **64**, 2839–2861, <https://doi.org/10.1175/JAS3980>.
- , G. Thompson, and V. Tatarskii, 2009: Impact of cloud microphysics on the development of trailing stratiform precipitation in a simulated squall line: Comparison of one- and two-moment schemes. *Mon. Wea. Rev.*, **137**, 991–1007, <https://doi.org/10.1175/2008MWR2556.1>.
- , M. R. Kumjian, C. P. Martinkus, O. P. Prat, and M. van Lier-Walqui, 2019: A general N -moment normalization method for deriving raindrop size distribution scaling relationships. *J. Appl. Meteor. Climatol.*, **58**, 247–267, <https://doi.org/10.1175/JAMC-D-18-0060.1>.
- , M. van Lier-Walqui, M. R. Kumjian, and O. P. Prat, 2020: A Bayesian approach for statistical–physical bulk parameterization of rain microphysics. Part I: Scheme description. *J. Atmos. Sci.*, **77**, 1019–1041, <https://doi.org/10.1175/JAS-D-19-0070.1>.
- Mosegaard, K., and A. Tarantola, 1995: Monte Carlo sampling of solutions to inverse problems. *J. Geophys. Res.*, **100**, 12 431–12 447, <https://doi.org/10.1029/94JB03097>.
- Posselt, D. J., and T. Vukicevic, 2010: Robust characterization of model physics uncertainty for simulations of deep moist convection. *Mon. Wea. Rev.*, **138**, 1513–1535, <https://doi.org/10.1175/2009MWR3094.1>.
- Sambridge, M., T. Bodin, K. Gallagher, and H. Tkalcic, 2013: Transdimensional inference in the geosciences. *Philos. Trans. Roy. Soc.*, **371A**, 20110547, <https://doi.org/10.1098/rsta.2011.0547>.
- Schrom, R. S., and M. R. Kumjian, 2018: Bulk-density representations of branched planar ice crystals: Errors in the polarimetric radar variables. *J. Appl. Meteor. Climatol.*, **57**, 333–346, <https://doi.org/10.1175/JAMC-D-17-0114.1>.
- Shipway, B. J., and A. A. Hill, 2012: Diagnosis of systematic differences between multiple parametrizations of warm rain microphysics using a kinematic framework. *Quart. J. Roy. Meteor. Soc.*, **138**, 2196–2211, <https://doi.org/10.1002/qj.1913>.
- Tridon, F., A. Battaglia, and D. Watters, 2017: Evaporation in action sensed by multiwavelength Doppler radars. *J. Geophys. Res. Atmos.*, **122**, 9379–9390, <https://doi.org/10.1002/2016JD025998>.
- van Lier-Walqui, M., T. Vukicevic, and D. J. Posselt, 2014: Linearization of microphysical parameterization uncertainty using multiplicative process perturbation parameters. *Mon. Wea. Rev.*, **142**, 401–413, <https://doi.org/10.1175/MWR-D-13-00076.1>.
- Vukicevic, T., and D. Posselt, 2008: Analysis of the impact of model nonlinearities in inverse problem solving. *J. Atmos. Sci.*, **65**, 2803–2823, <https://doi.org/10.1175/2008JAS2534.1>.
- Williams, C. R., 2016: Reflectivity and liquid water content vertical decomposition diagrams to diagnose vertical evolution of raindrop size distributions. *J. Atmos. Oceanic Technol.*, **33**, 579–595, <https://doi.org/10.1175/JTECH-D-15-0208.1>.